Review of ECOFRAM Terrestrial Draft Report

Comments from Dwayne Moore
The Cadmus Group
411 Roosevelt Ave., Suite 204
Ottawa Ontario K2A 3X9
Tel: (613) 761-1464
Fax: (613) 761-7653
moored@cyberus.ca

Overview

The ECOFRAM terrestrial draft report is the result of an ambitious attempt to update FIFRA assessment methods for estimating risks of pesticides to terrestrial biota. I believe that the approaches and methods described in the draft report represent a significant advance in the science of assessing the risks of pesticides to wildlife. The authors of this report also quite correctly recognize that data limitations and lack of understanding of how pesticides interact with the environment limit our abilities to predict effects or even to adequately characterize some sources of uncertainty. Even so, risk assessors have the capabilities to provide much more information on risks than is provided by the conservative quotients currently used to support decision making on pesticide acceptability and uses.

In general, the ECOFRAM report is scientifically sound and provides an excellent overview of state-of-the-art methods for estimating exposure, effects and risks. Most of my comments below are not major criticisms of the approaches and methods proposed. Instead, my comments are aimed at improving the text so that it can be more easily understood by risk assessors and risk managers. In places, I also suggest alternative methods that could be considered for assessing exposure, effects and risks of pesticides.

Instead of responding to the question listed in the charge to workshop panel members, I have chosen to organize my comments according to the EPA framework for ecological risk assessment. I found the questions listed in the charge constrained my ability to make the comments I wanted to make. In addition, organizing my comments by the framework should facilitate revisions to the report since the framework components correspond to the organization of the report.

Because of the time available to review this report and its length, I did not concern myself with editorial or minor comments. Also, I did not read the appendices in detail and thus will not offer comments on them.

Introduction

Most of the introduction chapter is an overview of probabilistic ecological risk assessment. This overview is useful in introducing readers to the steps involved in uncertainty analysis, and the advantages and limitations of conducting an uncertainty analysis. At times, however, I think the text is unnecessarily dense — most non-statisticians will not be able to comprehend or make use of the information presented in, for example, section 1.9.1. Also, it would be beneficial to the reader to introduce the conceptual steps involved in carrying out a probabilistic risk assessment earlier in the chapter. A figure would be useful to show these steps. The text gives the false impression that Monte Carlo simulation is nearly always going to be the best method for propagating uncertainties. I believe that other methods are often better suited to risk assessment problems, particularly where information is limited (as it always is). Finally, more care needs to be taken when describing and using basic terms in the text. Phrases such as "natural variation in the risk estimate" (pg. 1-11) "probability that a specific risk may occur" make no sense (risk is a function of probability and magnitude of effect) and are bound to confuse the reader. I suggest that formal definitions for terms such as risk (see Kaplan and Garrick 1981, Risk Analysis 1: 11-27 for an excellent definition), uncertainty, and variability be developed and included in a glossary. I elaborate on these comments below.

1-11 This section provides one of many typologies of sources of uncertainty in ecological risk assessment. This typology (and most others) is somewhat limited in that it does not acknowledge that probability can be thought of in several ways, and uncertainties are not always probabilistic in nature (uncertainty≠probability). Classical probability is based on the notion that past frequencies make good predictors. Relative frequencies are a way of characterizing uncertainties. It is not the only way, however, and ultimately may prove to be a rather impractical approach to characterizing uncertainties in ecological risk assessments where data are lacking on past performance of environmental parameters.

Another approach for characterizing uncertainties is to reason back from observed effects to cause, or more precisely, from evidence to hypothesis. Consider the example of an urn filled with black and white balls. A classical or frequentist approach would ask what is the chance that the next ball will be black, and base the answer on the number of black and white balls drawn previously from the urn. Bayesian or subjective probability would ask what is the ratio of black and white balls in the urn. Baye's theorum is essentially a formula that allows probabilities to be updated as new evidence becomes available. In the complete absence of information, Laplace's principle of indifference applies - all alternatives are equally likely. That is, prior to extracting any balls from the

urn, we start with the belief that the ratio of black and white balls in the urn is 1:1. This non-informative prior is then updated as balls are extracted. Eventually, the updated probability (the posterior) converges on the true ratio of black to white balls in the urn. Bayesian probabilities do not depend on a track record of past performance. As such, they are subjective and can be thought of as degrees of belief, rather than relative frequencies. Intuitively, this would seem to be a more useful paradigm for ecological risk assessments in which serious data gaps exist, but for which we may have much expert (*i.e.*, subjective) knowledge to draw upon.

Most Bayesians would argue that subjective probability is the only way to handle uncertainties in ecological risk assessments involving multiple stressors. There are, however, several types of uncertainties for which probability may be ill suited to describe. Consider the ambiguities inherent in our use of words to classify objects or describe systems. Suppose, for example, that one of the goals of the U.S. EPA is to protect sustainability of wildlife populations. The task for the risk assessors and managers would be to determine the probabilities of reductions in sustainability occurring given projections about pesticide usage and consequent effects in the region. Probability, however, does not capture the uncertainty that exists because the term sustainability lacks a precise definition. It is hard to estimate the probability of an event when the event covers a range of possibilities, some closer to the implied definition than others. Klir and Folger (1988. *Fuzzy Sets, Uncertainty, and Information.* Prentice-Hall, Englewood Cliffs, New Jersey) refer to this non-probabilistic type of uncertainty as fuzziness or vagueness. Another non-probabilistic source of uncertainty, nonspecificity or ambiguity, arises when the evidence is nonspecific and therefore uninformative as to the cause of the observed adverse effects. If, for example, we had evidence that songbird populations were declining in an area, we would not be able to determine whether this decline was due to pesticide use, habitat fragmentation, climate change or some other factor or combination of factors. In Klir and Folger's typology, the other non-probabilistic source of uncertainty (*i.e.*, confusion) arises when the very meaning of the evidence is unclear. In the end, probability only deals with dissonance (*i.e.*, pure conflict). A pesticide is or is not a carcinogen, or a pesticide may or may not kill a particular individual. Typically, some evidence supports one hypothesis, different evidence the other, and we are uncertain between the two. When the evidence is combined, probabilities may be assigned to each hypothesis. Thus, a bioassay may indicate that at a particular pesticide concentration, there is a 30% probability that a particular individual will die and a 70% probability that it will not die. With dissonance, a hypothesis cannot be both true and false at the same time, an outcome of Aristotle's law of the excluded middle.

The bottom line for this chapter is that I think it would be useful to educate readers as to the different schools of thought on probability, and recognize that non-probabilistic sources of uncertainty exist. Monte Carlo analysis (or Baye's theorum) cannot deal with these latter sources of uncertainty, but other methods exist that can (*e.g.*, fuzzy arithmetic).

1-12 In the absence of information, the text recommends that conservative models or assumptions be used to "represent reasonable worst case scenarios". This approach is fine for early tier assessments that are designed to screen out negligible risk scenarios (the conservative quotient approach). I do not feel it is an appropriate recommendation in probabilistic risk assessment, however, because the resulting risk estimates will be biased high. Worse, we will have no idea how biased the risk estimates are. This is often the criticism leveled at conservative quotients (*e.g.*, Cullen, A.C. 1994. Risk Anal. 14: 389-393; Bogen, K.T. 1994. Risk Anal. 379-381). Extending an analogy by Reckhow (1994. Ecol. Model. 72: 1-20), a forecast of "it will very likely rain" when rain is highly unlikely is not helpful; rather, we would like to know the true odds, and act according to our attitude toward risk. Thus, if the potential for severe effects exists, but uncertainty is high, risk managers may appropriately invoke conservatism by choosing to hedge their decisions away from the potentially serious effects by, for example, banning a pesticide use. By introducing conservatism into an assessment, assessors are in effect making policy decisions, decisions that should be made by risk managers and interested parties after being presented with an unbiased assessment. Use of expert judgment to estimate uncertainties about model structure or characterize input distributions is one way of dealing with some of the subjective uncertainties mentioned in this section. Meyer and Booker (1991. Eliciting and Analyzing Expert Judgment: A Practical Guide. Academic Press, NY) provides guidance on how to elicit information from experts.

The Lee and Wright reference should be 1994, not 1964.

1-14 The chapter by Mitchell Sharp in Morgan and Henrion (1990. Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK) provides an excellent discussion of how to select and parameterize input distributions.

1-15 The brief paragraph mentioning probability bounds analysis does not provide enough information to enlighten readers about the concept behind this methodology. Probability bounds analysis represents an uncertain input distribution with an entire class of probability distributions that conform with the available empirical information about the variable. Sometimes this class is very

small, and might be a single distribution when information is abundant. Other times, the class can be large, reflecting a poor state of knowledge about the variable. In Monte Carlo simulation, it is not possible to represent uncertainties about choice of distribution for an input variable, and it is difficult to represent uncertainties about parameter values (2^{nd} order Monte Carlo can do this, but it can be computationally difficult with large models and there is uncertainty about the choice of an appropriate 2^{nd} order error distribution). Ferson (1995, see reference in ECOFRAM draft) describes how to derive optimal probability bounds on the cumulative distribution functions in these classes for a variety of situations in which empirical information is limited. Using the numerical method developed by Williamson and Downs (1990. Int. J. Approx. Reasoning 4: 89-158), it is possible to compute bounds on the output cumulative distribution function when the input distributions are represented by probability bounds. In risk assessment, the use of probability bounds analysis reveals how much larger (or smaller) the probability of a result of a given magnitude might be. Probability bounds analysis offers a means for determining the reliability of risk estimates that is more comprehensive than what-if or interval approaches and yet computationally cheaper than Monte Carlo methods.

1-20 In cases where "the potential for adverse effects is high along with a high level of certainty", why would "further assessment ... be considered"? What would be gained from such an analysis? Mitigation decisions should be made quickly in such cases.

1-23 The discussion on PDFs and CDFs (section 1.9.1) will lose most readers. Given that the target audience for this document includes many assessors and managers not familiar with uncertainty analysis, I suggest that a more user-friendly description of normal and lognormal distributions be used here. Readers do not need to know the equations, just the concepts and assumptions behind the distributions. Figures could be used to show distribution shapes for both PDFs and CDFs.

Section 1.9.1 presents a lot of detail on normal and lognormal distributions, and no information on other distributions likely to be encountered in ecological risk assessments (*e.g.*, beta, logistic, Weibull, chi square, Poisson, binomial). The concept of probability mass functions should also be introduced here.

1-26 Original references should be given where the text describes methods used to induce dependencies (*e.g.*, Pearson correlations) in Monte Carlo simulation. These include Scheuer and Stoller (1962. Technometrics 4: 278-281), Iman and Conover (1982. Communications in Statistics B11: 311-334), and Nelson (1986. Communications in Statistics A15: 3277-3285).

1-28 Chapter 1 gives the impression that Monte Carlo simulation will be the only method used to develop exposure and risk distributions. There are many other methods that may be used to propagate uncertainties including Baye's Theorum (briefly introduced, but not referred to any further in the document), probability bounds analysis, first order error analysis (for simple models), fuzzy arithmetic, and others. Bayesian methods, in particular, have a long history of use and are amenable to quantitative cost-benefits analyses that can be used to aid the decision-making process (*e.g.*, Dakins *et al.* 1994. Environ. Toxicol. Chem. 13: 1907-1915). A more thorough discussion of uncertainty propagation techniques is warranted.

Problem Formulation

This chapter and the expanded text in Appendix B was useful and well written. My only quibble was with the adoption of the "assessment endpoints" and other EPA framework terms. Over the last few years, I have found that these terms and their definitions often create much confusion. I much prefer the "risk scenarios" terms and definitions developed by Kaplan and Garrick (see reference cited above). I realize, however, that the Framework terms are now part of the ecological risk assessment paradigm and that to adopt a different terminology for FIFRA assessments would create many difficulties within the Agency.

2-3 The list of risk management questions is well thought out. Assessors would do well to use this list to guide discussions with risk managers and to plan risk assessments.

2-8 Some of the community and system values endpoints are vague. Taxonomic diversity, for example, is a function of evenness and richness (better to use the two terms separately). The term "functional diversity" is not one I have encountered in the ecological literature. "Compositional integrity" is a vague term.

2-12 The section on defining the ecosystem at risk is a useful one and provides insights into some of the spatial and temporal scale issues that one must consider. One possible outcome of a pesticide application, *i.e.*, that unexposed populations can be adversely affected (action at a distance, see Spromberg et al. 1998. Environ. Toxicol. Chem. 17: 1640-1649), was, however, overlooked. Some discussion of this possible outcome and metapopulation dynamics in general is warranted.

Exposure Assessment

I found that this chapter and the expanded discussion in Appendix C to be quite useful, well written and provided a wealth of information on exposure models and the variables to be considered in estimating exposure. Particularly useful was the section describing environmental databases. The only major recommendation I would make would be to provide more insight on what distributions would be appropriate (with accompanying rationales) for different input variables. Otherwise, novice users may simply default to distributions that provide the best goodness-of-fit without first considering theoretical plausibility.

3-26 Moore *et al.* (1997. Environ. Toxicol. Chem. 16: 1042-1050; 1999. Environ. Toxicol. Chem. 18(12)) describe two case studies showing how uncertainty in estimation of metabolic rates, gross energy of prey items, assimilation efficiencies of prey items, and mixed diets may be incorporated to estimate exposure distributions for mink.

3-27 It would be useful to specify what the distributions would be for assimilation efficiency, gross energy and metabolic rate. The beta distribution is a good choice for assimilation efficiency because it is continuous and can be scaled from 0 to 1 (no need for truncation).

3-29 Because proportions of animal and plant matter in the diet scale from 0 to 1, I would have chosen a beta distribution for these dietary variables, rather than normal distributions (the latter require truncation at zero and one). Where there are more than two prey items in the diet, negative correlations between proportion prey items in the diet would need to be introduced (*i.e.*, changes in proportion consumption of one prey item would presumably produce opposite changes in proportional consumption of other prey items).

3-31 Means and standard deviations are efficient statistical estimators only when the underlying distribution is normal. Because many input variables will not have underlying normal distributions, it may not be advisable to use means and standard deviations to “develop hypothetical distributions” in refined assessments.

3-33 If avoidance is a function of dietary exposure, then this dependency should be induced in the Monte Carlo analysis (easily done in Crystal Ball).

3-38 In the second paragraph of section 3.3.9, it would be useful to explain the consequences of repeated instantiation of variables in model equations.

The standard deviation in the last line is too small to be realistic.

3-39 Why would “truncated” normal distributions be used for body weight?

3-42 Is there sufficient data or knowledge to specify appropriate distributions for partition coefficients, Henry’s law constant, abiotic degradation rates, etc?

3-72 The section on computerized exposure models is generally well written, except that there is little mention of whether the models (PRZM3, EXAMS, etc) have undergone verification and validation studies. It would be very useful to know the predictive capabilities of these models. Also, in this section it would be helpful to describe the different kinds of statistics that can be used to measure model performance. Potential techniques include lumped measures of average model goodness-of-fit, correlation measures, parametric and non-parametric statistical tests, spatial analysis of goodness-of-fit (*e.g.*, kriging), and Bayesian measures of estimation error. Errors due to model structure can be included in uncertainty analyses (a topic not discussed in this chapter).

In several places in this section, there is a call to develop a “comprehensive terrestrial exposure model”. Presumably, such a model would be massive and have perhaps 100s of input variables. Before embarking on such a task, it would be useful to read Reckhow’s (1994. Ecol Model. 72: 1-20) article cautioning against the use of large, mechanistic fate and transport models (*e.g.*, WASP4 and EXAMS) in regulatory decision making because of their poor track record in producing reasonably accurate predictions. In an excellent review paper, Beck (1987. Water Resour. Res. 23: 1393-1442) concluded that “most of the evidence suggests that the current models of water quality, in particular, the larger models, are easily capable of generating predictions to which little confidence would be attached”. I doubt the situation is any better with terrestrial models.

3-83 The section on environmental databases was very helpful.

3-98 The chi-square statistic is not the only goodness-of-fit statistic available. Note that because the residual term is squared, there will be a tendency for outliers to have a profound effect on the estimated goodness-of-fit according to this statistic.

3-99 Seiler and Alvarez (1995) were highly critical of the use of triangular and uniform distributions in ecological risk assessment because the distributions simply do not describe environmental phenomena. The explanation on this page that their criticism was “due in part to discontinuities in those distributions” is obtuse and misses the point for most readers.

3-99 The use of expert judgment to derive input distributions when data are limited should be discussed in this section.

Effects Assessment

Although this chapter contains much useful information, I found it to be the weakest chapter in the report. In my opinion, the chapter was flawed because the authors did not think much "outside the box". Too often, the methods recommended were constrained by the types of toxicity data currently required under FIFRA. While such constraints are real and need to be recognized, it should not prevent the authors from recommending methods that could be used should more comprehensive data sets be available. Comprehensive data sets (*e.g.*, data sets that include chronic studies with non-lethal endpoints, mammal studies, invertebrate studies, etc) are sometimes available for high use pesticides (*e.g.*, the triazines) and could be required in the future for pesticides undergoing level 4 assessments.

Based on the above considerations, I believe that the chapter should begin by describing a comprehensive statistical framework for the analysis of toxicity data. For deriving dose-response relationships, a useful framework is the Generalized Linear Model (GLiM) framework recently described by Bailer and Oris (1997. Environ. Toxicol. Chem. 16: 1554-1559). The GLiM framework includes two components: (i) the probability distribution of the response variable, and (ii) a link function or transformation that effectively linearizes the concentration-response or dose-response relationship to facilitate linear regression analysis. Kerr and Meador (1996. Environ. Toxicol. Chem. 15: 395-401) and Bailer and Oris (1997) describe this framework in detail. In the GLiM framework, the probability distributions are the binomial, Poisson and normal (or Gaussian) distributions for quantal, count and continuous responses, respectively. Probit or logit link functions are used for quantal responses, and the log link function for other responses. The GLiM framework can thus deal with a variety of wildlife responses (*e.g.*, mortality, number of young, biomass, etc). By adding a quadratic term to the regression equation (*i.e.*, $y = a + bx + cx^2$), the framework is also capable of dealing with unusual relationships such as stimulation at low dose. The chapter would be much improved by adopting and describing a comprehensive statistical framework such as the GLiM framework for analysis of toxicity data, rather than describing the much more limited probit model in detail.

The second flaw I found in this chapter was the almost total absence of information on methods to develop dose-response relationships for mammals (mammals were covered in the exposure chapter), reptiles, amphibians, invertebrates, and terrestrial plants. I realize that the "standard" FIFRA data set does not include these organisms, but surely data are sometimes available for organisms other than birds with high use existing pesticides, or could be required in level 4 assessments. Also, much stronger

recommendations could be made to expand “standard” FIFRA data sets to include organisms other than birds. Finally, some discussion could be provided on existing QSARs that have been developed for extrapolations of bird toxicity test results to other organisms such as mammals. Uncertainties in the QSAR relationships can be explicitly accounted for in probabilistic risk assessments of pesticide effects to, for example, mammals.

4-4 Much of the material on this page and the top of page 4-5 repeats earlier material.

4-14 In section 4.1.4 and elsewhere, the text needs to emphasize that the probit model is only appropriate for quantal responses such as mortality. Also, in section 4.1.4, it would be useful to show some of the more important model equations (as was done in chapter 3) or at least provide references to the literature so that assessors can learn more about the models.

4-17 A variety of threshold models (*e.g.*, segmented models) can explicitly derive a threshold (with confidence or fiducial limits). Such models should be briefly described here.

4-18 At the top of this page, it is not clear what approach is being referred to for creating slope estimates from existing data. Were the authors referring to meta analysis or some other approach? Whatever the approach, the steps involved need to be briefly described, otherwise this material is of little use to the reader.

4-19 No guidance or methods on how to do time to event or dose to event analyses is provided in the text. Also, references to the literature should be provided here. This comment applies to much of this chapter — lots of issues and methods are mentioned, but little guidance is provided to help assessors.

4-20 Is the OECD developing test guidelines for mammals, invertebrates, plants, reptiles and amphibians? If not, a recommendation to develop such test guidelines should be made.

4-21 I am mystified why the terrestrial workgroup decided against recommending for a redesign of the current avian reproduction study. Surely there is sufficient literature available to have convinced the workgroup that NOECs and LOECs are seriously flawed endpoints for toxicity testing. To facilitate derivation of dose-response relationships, the avian reproduction study should be redesigned to have more treatments and fewer replicates. This would not add to the costs of carrying out such studies and statistical methods are already well developed for deriving dose-response relationships. The argument that a study redesign is not

justified because there are so many uncertainties involved in extrapolating from the laboratory to the field or between species is facile. Such narrow thinking would quickly lead to stagnation of methods development if it is used every time uncertainties are encountered in ecological risk assessment. We make improvements as we can.

4-24 Section 4.3.1 notes the importance of considering indirect effects and the current limitations of the pesticide registration process for addressing such effects, but stops short of making any recommendations for improving the situation. What should be done in the future to give assessors the capability to address indirect effects?

I have no idea what the first sentence in section 4.3.2 means.

4-27 The method proposed for dealing with sublethal effects in risk assessment (lines 7-9) is inadequately described. Are the authors referring to the use of mixture models to combine lethal and non-lethal responses? My understanding is that such models have little history of use in toxicology. Guidance is required here, because few assessors will be familiar with the use of mixture models.

4-29 This section and elsewhere seems to focus almost exclusively on the probit model, because “to use results based on the probit model in some alternative model ... would not necessarily be straightforward”. Maximum likelihood procedures are, however, readily available for fitting all sorts of models to raw data. Thus, there is no need to transform the probit model parameters (*i.e.*, mean, standard deviation) to the parameters required in other models. The reason for including so much detail on probit models at the expense of considering other models in this chapter continues to elude me.

4-37 Confidence limits at the 5 or 10% response levels are not necessarily substantially wider than at the 50% response value. If several treatments span the no and low toxic effects portion of the dose-response curve, the confidence limits will be similar to or narrower than those that occur at higher response levels.

4-38 Section 4.4.3.2 is really obtuse. What is the point of this section?

4-42 I am not sure why methods such as hierarchical Monte Carlo, Bayesian theory, and Monte Carlo “coverage” experiments are described here. Without a better introduction to the methods and the reasons for using them, this section will lose most readers.

4-44 Terms such as marginal and joint distributions need to be defined. Using hierarchical Monte Carlo to derive distributions on probit parameters seems a particularly obtuse way for dealing with parameter uncertainties. Kerr and Meador (1996) (and many statistics textbooks) provide a much more straightforward algebraic approach for estimating confidence or fiducial limits for the probit model. The only information required are the parameter estimates, their standard errors, and the correlation between the parameters — all of these are standard outputs from just about any statistical package that does regression analysis.

4-45 That no distribution is assumed for an independent variable is only true in type I regression.

4-56 Mineau *et al.* (1996) found that the mean allometric scaling factor for chemical toxicity to birds was 1.15 and ranged as high as 1.55. Sample and Opresko (1996. Toxicological benchmarks for wildlife: 1996 revision. ES/ER/TM-86/R3. Oak Ridge National Laboratory, Oak Ridge, TN, USA), however, state that an allometric scaling factor of one is still appropriate for birds. This is because, for the majority of chemicals in the Mineau *et al.*, study factors were not significantly different from one. As a result,

$$ED_p^w = ED_p^t \left(\frac{bw^t}{bw^w} \right)^0 \quad (1)$$

where ED is the effects dose, p is proportion affected, w is avian wildlife species, t is avian test species, and bw is body weight. Therefore, they argue that toxicity test results normalized to body weight (*i.e.*, μg/kg body weight/day) can be used to estimate effects doses for other avian wildlife species without use of an allometric scaling factor.

4-59 The approach described for deriving extrapolation factors seems reasonable, and should produce more defensible factors than is the case with the generic factors currently used. Note, however, that the methods described are designed to predict the 5th percentile LD_{50} -- thus, there is a 95% probability that the focal species could be less sensitive (a conservative bias). On the other hand, the LD_{50} is a very serious adverse effect for the 5th percentile species. Should the aim be to protect 95% of individuals of sensitive species? Since these issues are policy decisions better handled during the risk management phase, an alternative approach would be to use the input distributions in a Monte Carlo analysis to derive distributions for the focal species LD_{50}s and other responses. These outputs could then be combined with exposure distributions to develop

risk curves.

Risk Assessment Methodology

In general, this chapter was well written, although I do have several criticisms. First, as I have noted before, the only uncertainty propagation technique considered is Monte Carlo simulation. Other techniques should be described and illustrated in a case study. Second, the chapter did not provide much information on how to communicate the results of the analyses to risk managers and stakeholders, nor on how to use the information in decision making. Chapter 3 in Warren-Hicks and Moore (1998. Uncertainty Analysis in Ecological Risk Assessment. SETAC Press, Pensacola, FL) provides a general overview for communicating uncertainty analysis results to different audiences. Revisiting the risk manager questions in chapter 2 of this report would also be a useful exercise to help guide assessors in communicating their results to decision makers. Decision analytic approaches can be used as a quantitative tool for optimizing decisions in the face of uncertainties (see recent debate series in Human and Ecological Risk Assessment 5(2)).

5-7 I think method 4 (distribution-based quotients) could be eliminated. If exposure and effects distributions are available, why not just integrate them to produce risk curves (a computationally easy exercise) as specified in method 5. Risk curves are much more informative than are probabilistic quotients.

5-8 The estimated 90^{th} and 95^{th} percentiles for exposure exceed the 100% exposure value from the cumulative exposure distribution. I realize that this occurred because a lognormal distribution was fit to a hypothetical data set, but nevertheless this could be confusing to readers.

5-9 The use of LD_{50}s in the risk examples means that the case studies are estimating probabilities of very serious adverse effects. Thus even if the probability of exceeding the LD_{50} are low, this will not necessarily mean that risks are low.

5-11 I have my doubts about some of the limitations and advantages listed for the quotient approach. For example, the method could account for space and time simply by calculating quotients for different areas or at different times. Given that quotients are generally calculated with varying mixtures of conservatism, I am not sure I agree that quotients can be used to compare risks among alternative chemicals.

5-14 The figure gives the impression that the toxicity threshold (NOEC) exceeds the estimated LD_{50}. The caption needs to clarify that these toxicity endpoints came

from different studies.

5-15 In my opinion, the Margin of Safety (MOS) approach is a poor way of summarizing and communicating the results of a probabilistic risk assessment. The choice of the 90th percentile exposure value and 10th percentile effects value is purely arbitrary, and is essentially making a policy decision about how conservative the assessment should be. If there is sufficient information to estimate an MOS, then there is sufficient information to derive a risk curve. A risk curve communicates much more information to the risk manager (*e.g.*, probability that 10% of individuals will die) than does the MOS approach, and does not make arbitrary policy decisions. I strongly suggest deleting the MOS as a recommended approach in this section.

5-16 I would not have expected the shapes of the exposure and effects distributions to be linear as shown in Figure 5.5.1 to 5.5.3. More typical distributions should be used in these examples (*e.g.*, lognormal exposure CDF, log-logistic effects distribution).

5-34 Appendix D1 did not discuss the issue of how mortality can be generated directly from the dose-response model as is stated in the text at the top of this page. I am curious as to how many individuals are required before one can ignore the use of the random mortality component.

5-36 Case studies should be provided to illustrate the use of survivorship models and mechanistic population models.

Levels of Refinement for the Assessment Process

I liked this chapter and agree with the suggested levels of refinement. I have only one minor comment. Table 6.2-1 gives the impression that an LD_{50} will be the denominator in the quotient. Since level 1 is being used to screen out negligible risk scenarios, I suggest that the LD_{50} needs to be converted to a no or low toxic effects level (*e.g.*, LD_5) before quotients are calculated.

Recommendations

I agreed with most of the recommendations listed in this chapter and with the self evaluation of how the workgroup fulfilled its charge. The only recommendation I would add would be to encourage that standard FIFRA data sets be expanded to include toxicity test results for mammals and other terrestrial organisms.